



XXVIII SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2024

CARACTERIZAÇÃO TOPOLOGICA DE REDES SEMÂNTICAS BASEADAS EM LIVROS DIDÁTICOS DE MATEMÁTICA

Matheus Oliveira dos Santos¹; Marcos Grilo²

1. Bolsista CNPq, Graduando em Licenciatura em Matemática, Universidade Estadual de Feira de Santana, e-mail: matheusods10@gmail.com;
2. Orientador, Departamento de Ciências Exatas, Universidade Estadual de Feira de Santana, e-mail: grilo@uefs.br

PALAVRAS-CHAVE: Redes Semânticas; Livro Didático; Caracterização topológica.

INTRODUÇÃO

Pesquisas com redes semânticas são aplicadas na análise de discursos escritos (Caldeira et al., 2006) e orais (Teixeira et al., 2010), na difusão do conhecimento por títulos científicos (Fadigas et al., 2009; Pereira et al., 2011) e na identificação de temas em trabalhos de conclusão de curso (Santos & Grilo, 2020). Neste plano de trabalho, objetivamos caracterizar a estrutura topológica de redes semânticas baseadas em sentenças de livros didáticos de Matemática.

METODOLOGIA

Para a construção da rede semântica de cliques, utilizamos como fonte de dados o capítulo “Matrizes e Determinantes” do livro Matemática: Contexto & Aplicações, volume 2, 3^a edição, de 2016, do autor Luiz Roberto Dante. O livro escolhido é anterior à Base Nacional Curricular Comum (BNCC). “Matrizes e Determinantes” é um conteúdo que integra os currículos de Matemática do Ensino Médio e dos Cursos de Engenharia de Computação e Licenciatura em Matemática no Brasil. Aplicamos um tratamento manual, seguido de um tratamento computacional por meio do TEX-2-NET (SANTOS e GRILLO, 2023).

No tratamento manual, figuras, esquemas, gráficos, tabelas, exercícios e o preâmbulo do LaTeX foram excluídos do texto minerado pelo Mathpix. Estabelecemos regras para o tratamento manual que visaram preparar o arquivo de texto para que expressões, variáveis e símbolos matemáticos sejam incluídos na rede semântica. A critério do pesquisador, considerando o contexto do capítulo do livro didático escolhido, as regras específicas podem ser definidas a fim de melhorar a qualidade dos dados. O método de construção das redes semânticas de cliques segue as técnicas propostas por Caldeira, et al. (2006), Fadigas et al. (2009), e Santos e Grilo (2023).

RESULTADOS

Uma rede G esparsa, não dirigida, sem arestas múltiplas, sem loops e conectada, é dita mundo pequeno, pelo modelo de Watts & Strogatz (1998), se o caminho mínimo médio L é baixo e o coeficiente de aglomeração médio C é alto quando comparados a

uma rede aleatória equivalente. Para o cálculo desses índices, usamos a biblioteca de código livre Networkx (Hagberg *et al.*, 2008). A Tabela 1 apresenta índices para a rede semântica de clique construída a partir do capítulo “Matrizes e Determinantes” do livro Matemática: Contexto & Aplicações, volume 2, 3^a edição, de 2016, do autor Luiz Roberto Dante, denominada de rede G. Os índices foram calculados pela biblioteca Networkx. Trata-se de uma rede finita, não-vazia, rotulada, de 1-modo, sem arestas múltiplas, sem laços, não dirigida e ponderada. Para avaliar o efeito mundo-pequeno (WATTS e STROGATZ, 1998) na maior componente da rede G, denominada de rede GMC, geramos uma rede aleatória equivalente denominada de GMC_{random}, usando o software Pajek.

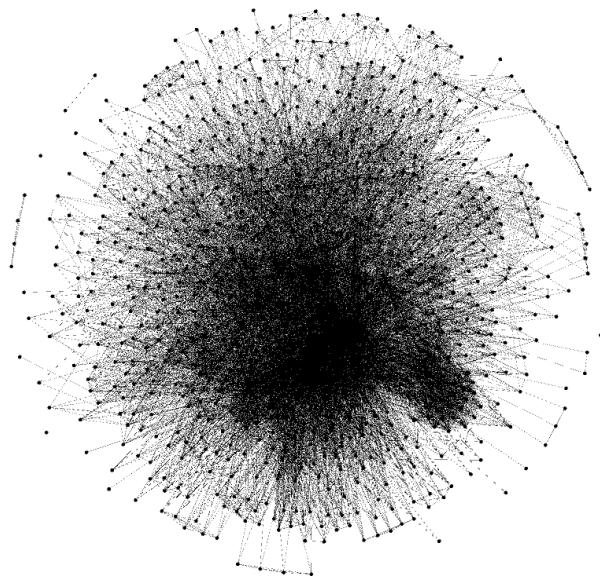
Tabela 1. Índices das redes G, GMC, GMC_{random}.

Rede	G	GMC	GMC _{random}
Número de Vértices	787	777	777
Número de Arestas	12044	12036	12051
Densidade	0,0389	0,0399	0,0399
Grau Médio	30,6073	30,9806	31,0193
Maior Componente	777	777	777
Caminho Mínimo Médio	2,3181	2,3182	2,2368
Coeficiente de Aglomeração Médio	0,7712	0,7621	0,0398
Diâmetro	6	6	3

Fonte: autores

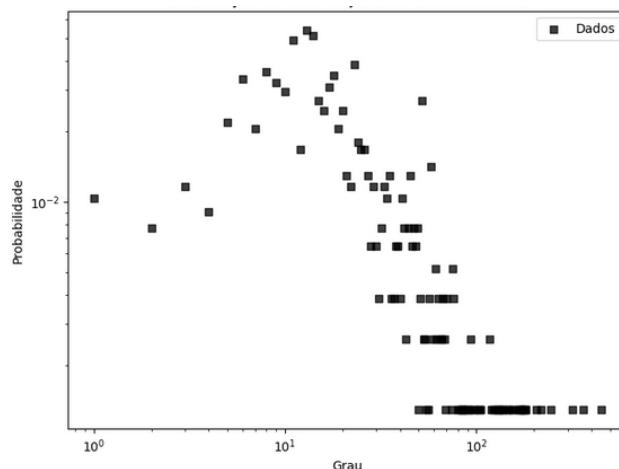
Os valores dos índices da rede GMC quando comparados com a rede aleatória equivalente GMC_{random} permite-nos afirmar que a rede G possui uma estrutura mundo-pequeno conforme proposto por Watts e Strogatz (1998). Barabási e Albert (1999) propuseram o modelo de rede livre de escala, isto é, redes cuja distribuição de graus obedece a um regime de lei de potência. Clauset *et al.* (2009) desenvolveram um conjunto de técnicas estatísticas para avaliar se a hipótese de lei de potência é plausível para os dados observados. Para verificar a hipótese livre de escala conforme Clauset *et al.*(2009), usamos o software livre Octave. Calculamos o parâmetro p-value, cujo valor foi igual a zero. Desta forma, segundo Clauset *et al.* (2009), a hipótese de ser livre de escala não é plausível para a rede G. A Figura 1 apresenta uma visualização da rede G, e a Figura 2 apresenta a distribuição de graus da rede no ajuste determinado pelo algoritmo de Clauset *et al.* (2009).

Figura 1. Visualização da rede semântica GMC.



Fonte: autores

Figura 2. Distribuição de graus da rede GMC.



Fonte: autores

Isso evidencia uma tendência de comportamento, uma vez que o modelo mundo-pequeno foi identificado em redes semânticas de cliques construídas a partir de fontes de dados diversas: discursos escritos (CALDEIRA et al., 2006); discursos orais (TEIXEIRA et al., 2010); títulos de artigos científicos (FADIGAS et al., 2009; PEREIRA et al., 2011); títulos de trabalhos de conclusão de curso (SANTOS e GRILLO, 2020); títulos de dissertações de mestrados (FADIGAS et al., 2020). Ao mesmo tempo, a aplicação do método de Clauset et al. (2009) apontou a não plausibilidade do comportamento de livre de escala nessas redes semânticas de cliques.

CONSIDERAÇÕES FINAIS

A caracterização da rede semântica construída do capítulo "Matrizes e Determinantes" do livro didático de Luiz Roberto Dante revelou que se trata de uma rede

mundo-pequeno no modelo do Watts e Strogatz (1998) e a plausibilidade de ser livre de escala não foi comprovada. Algumas tarefas do tratamento manual foi possível automatizar por meio do refinamento da ferramenta computacional TEX-2-NET, como a remoção dos comandos de formatação do LaTeX.

REFERÊNCIAS

- BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, n. 286, p. 509–512, 1999.
- CALDEIRA, S. M. G.; LOBÃO, T. C. P.; ANDRADE, R. F. S.; NEME, A.; MIRANDA, J. G. V. The network of concepts in written texts. *The European Physical Journal B*, v. 49, p. 523-529, 2006.
- CLAUSET, Aaron; SHALIZI, Cosma Rohilla; NEWMAN, Mark EJ. Power-law distributions in empirical data. *SIAM review*, v. 51, n. 4, p. 661-703, 2009.
- FADIGAS, I. S.; HENRIQUE, T.; DE SENNA, V.; MORET, M. A.; PEREIRA, H. B. B. Análise de redes semânticas baseada em títulos de artigos de periódicos científicos: o caso dos periódicos em educação matemática. *Educação Matemática Pesquisa*, v. 11, n. 1, p. 142-164, 2009. Acesso em: 08 ago. 2024.
- HAGBERG, A.A.; SCHULT, D.A.; SWART, P. J. Exploring network structure, dynamics, and function using NetworkX. In: VAROQUAUX, G  el; VAUGHT, Travis; MILLMAN, Jarrod (Eds.). *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Pasadena, CA, USA, p. 11–15, ago. 2008.
- PEREIRA, H. B. B. et al. Semantic networks based on titles of scientific papers. *Physica A: Statistical Mechanics and its Applications*, v. 390, n. 6, p. 1192–1197, 15 mar. 2011.
- S  , P. F. d.; J  NIOR, A. L. L. d. Livros did  ticos de matem  tica do 9  o ano do ensino fundamental: uma an  lise ´a luz da matriz de refer  ncia da prova SAEB.
- SANTOS, M. O.; GRIFO, M. Desenvolvimento de Ferramenta Computacional para Conversão de Representações Simbólicas em Livros Didáticos no Formato OCR PDF para TXT. *Anais dos Seminários de Iniciação Científica*, UEFS, 2023.
- SANTOS, V. C. DOS; GRIFO, M. Identificação de temáticas de trabalhos de conclusão de curso por meio de redes semânticas. *Revista Paranaense de Educação Matemática*; Vol. 9, No 20 (2020).
- TEIXEIRA, G. M. et al. Complex Semantic Network. *International Journal Modern Physics C*, v. 21, n. 3, p. 333–347, 2010.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. *Nature*, v. 393, n. 6684, p. 440–442, 1 jun. 1998.